

DOCUMENT RESUME

ED 177 212

TM 009 733

AUTHOR Doss, David A.; Ligon, Glynn D.
 TITLE The BRC: A Practical, Economical, Reliable, and Valid Measure of Student Classroom Behavior. Publication No. 78.78.
 INSTITUTION Austin Independent School District, Tex. Office of Research and Evaluation.
 PUB DATE Apr 79
 NOTE 20p.; Paper presented at the Annual Meeting of the American Educational Research Association (63rd, San Francisco, California, April 8-12, 1979)
 EDRS PRICE MF01/PC01 Plus Postage.
 DESCRIPTORS *Behavior Rating Scales; Compensatory Education Programs; *Counseling Effectiveness; Elementary Education; *Elementary School Counseling; Factor Structure; Misbehavior; Parent Participaticn; Program Evaluation; *Student Behavior; Student Motivation; Test Construction; Test Reliability; Test Reviews; Test Validity
 IDENTIFIERS Austin School District TX; *Behavior Rating Checklist; Devereux Elementary School Behavior Rating Scale; Elementary Secondary Education Act Title I; Texas (Austin)

ABSTRACT

The Behavior Fating Checklist (BRC) was developed in the Austin Independent School District to evaluate the counseling component in the Elementary Secondary Education Act Title I compensatory education programs for students in kindergarten through grade 5. The BRC was developed by: defining the behavior on which the counseling activities focused; studying other behavior rating scales; drafting items for pilot testing; pilot testing; analyzing the pilot tests; conferring with program staff and parent advisory committees; and finalizing the instrument. The final instrument was comprised of ten items to be rated on a nine-point scale. Five items measured students' on-task readiness, willingness, and ability to participate; and five items measured students' in-class discipline or disruptive behavior. Two evaluations of the BRC were conducted: (1) an item-by-item analysis to check factor structure and reliability and (2) a check of the validity of BRC. Conclusions were reached indicating that BRC is practical, economical, reliable, and valid.
 (MH)

 * Reproductions supplied by EDRS are the best that can be made *
 * from the original document. *

ED177212

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

THE BRC: A PRACTICAL, ECONOMICAL, RELIABLE, AND VALID
MEASURE OF STUDENT CLASSROOM BEHAVIOR

A paper presented at the annual meeting of the
American Educational Research Association

April, 1979

Authors: David A. Doss
Glynn D. Ligon

Publication No. 78.78

© 1979, David A. Doss and Glynn D. Ligon
Office of Research and Evaluation
Austin Independent School District
Austin, Texas 78752

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

David Doss
Glynn Ligon

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

TM 009 733

THE BRC: A PRACTICAL, ECONOMICAL, RELIABLE, AND VALID MEASURE
OF STUDENT CLASSROOM BEHAVIOR

The Behavior Rating Checklist (BRC) allows a teacher to quickly rate a student's behavior on ten items which divide into two scales - 1) Ready, Willing, and Able and 2) Descriptive Classroom Behavior. Reliability and validity studies indicate that the BRC is a dependable measure of a student's behavior in the two areas cited.

Austin's ESEA Title I Program, like many other compensatory education programs, has made a sizeable commitment of resources to a counseling and guidance component. Prior to 1975, increases in student attendance were the focus of the evaluation. Dissatisfaction with this criterion of success was evident within both the counseling staff and the evaluation staff. Something more directly a result of the counselor's efforts was needed to bring the evaluation and the programmatic activities into a closer relationship.

In the fall of 1975, a decision was made to measure changes in students' behavior as a result of the counseling services delivered. The Devereux Elementary Behavior Rating Scale (DEBES) was selected as the most appropriate of the behavior rating instruments available. Resources were not provided for direct observation of student behavior by evaluation staff.

Using the DEBRS was a mistake. For the sake of brevity, here is a list of some of the reasons the DEBRS had only a one-year engagement with the Title I Program.

1. The 47 items required teachers to contemplate from 10-45 minutes per student to provide their ratings. The one-third

sample used demanded up to several hours for some teachers to complete the ratings.

2. Some teachers questioned the appropriateness of their making judgements about student behavior in areas not directly observable in the classroom.
3. Counselors felt that many of the items referenced areas of behavior that were beyond the scope of their efforts.
4. The Title I Program's Parent Advisory Council (PAC) challenged the qualifications and objectiveness of classroom teachers for rating their children's behavior in some areas.
5. A comparison of students to the norm group representative of a very dissimilar family background showed Austin's Title I Program students to be much less well behaved. Parental reaction to this obscured the real purposes of the ratings - that being to measure change during the school year.

For the 1976-1977 school year, a new approach was needed. An examination of the problems encountered with the DEBRS and the purposes for rating student behavior resulted in the following criteria's being set for a replacement instrument.

The replacement behavior rating instrument must...

1. Be directly related to the student behavior on which the counseling activities focus,
2. Call for teachers to rate behavior which is within their field of observation and within their qualifications,
3. Require a minimum amount of time to complete,

4. Be oriented to the measurement of teachers' ratings of student behavior at different times for the purpose of investigating change rather than for the purpose of the comparison of Austin students to a normative sample, and
5. Produce a scale score rather than narrative descriptions.

A review of the instruments cited in Buros' Metal Measurements Yearbook failed to produce an acceptable substitute for the DEBRS. Criteria 1, 2, and 3 were especially absent in instruments already developed. Additionally, reliability and validity data were scarce on most behavior-rating scales.

The only alternative was to develop a rating scale appropriate for Austin's puposes.

The development of a scale proceeded through the following steps.

Step 1: Defining the behavior on which the counseling activities focused.

Through a study of the proposed activities of the counseling component, two areas of classroom behavior were identified as being targets of the counseling activities.

1. Student on-task and productive use of instructional time.
2. Student in-class discipline.

Step 2: Reviewing other behavior-rating instruments for study of scales and/or items which appeared to focus on the target behaviors.

Step 3: Drafting items for pilot testing.

Step 4: Pilot testing.

Six teachers in one of the program schools critiqued

the instrument after using it to rate some of their students.

Step 5: Analyzing the pilot test.

An item analysis was conducted using program TESTAT from the VSTAT statistical package, University of Texas at Austin. This analysis showed acceptable correlation of each item to a total score and excellent spread of the ratings across the full range of the scale for each item.

Step 6: Conferring with program staff and the Parent Advisory Committee.

Although the parents still were uneasy about the biases of some teachers and about the misuse of the results to show their children's behavior as poor, they agreed that the new instrument was a definite improvement over the DEBRS for Austin's purposes.

Step 7: Finalizing the instrument.

The final format of the instrument contains 10 items to be rated on a nine-point scale:

There has been
no evidence of
this behavior.

This behavior
occurs occa-
sionally.

This behavior
is frequent
and typical.

1 2 3 4 5 6 7 8 9

The items can be found in Figure 5. Five items were designed to measure student on-task and productive use of instructional time, and five others were designed to measure student in-class discipline.

The Behavior Rating Checklist (BRC) was used in the evaluation of the Counseling and Guidance Component during the 1976-77 school year and then again during 1977-78 and 1978-79. The analyses for the evaluation were done on a item-by-item basis during the first two years; however, a series of analyses have recently been done to test the feasibility of combining the items into two scales to measure the two behaviors which were thought to be measured when the instrument was developed. This paper reports on the results of those analyses.

Study I

The first study was done to check the factor structure and reliability of the BRC. The first study was specifically designed

- a. to verify that the items did indeed measure two areas of classroom behavior,
- b. to assign BRC items to scales to measure those behaviors, and
- c. to evaluate the reliability of those measures.

Subjects

Behavior ratings were obtained for a random sample of students in the 19 Title I schools with a Guidance and Counseling Program. Altogether 1,762 students in grades K-5 were rated in October, 1977, and April, 1978, by their classroom teachers. The number of students receiving ratings by grade can be found in Figure 1.

Analyses

In order to verify that the BRC taps two areas of classroom behavior and that the instrument measures the same thing pre and post at each grade level a series of factor analyses was done. First a noniterative,

principal components analysis with no rotation was done to determine the number of factors. Then principal axes factor analyses with orthogonal and oblique rotations were done for the pre- and postratings. The factor analyses were done using the SPSS Factor Analysis program on the University of Texas at Austin's CDC 6000 computer system. The orthogonal rotation was a Varimax rotation with Kaiser normalization. The oblique rotations were direct oblimin rotations with the delta value set at 0 (the default value).

After the items were placed on scales based on their factor loadings, the scales were scored and their internal consistency and test-retest reliabilities were evaluated. Cronbach alpha coefficients of internal consistency were calculated using the TESTAT program in Veldman's (1971) EDSTAT-V package of computer programs. Test-retest reliabilities were checked by correlating pre- (October) and postratings (April) using program Pearson Corr in the SPSS package.

Results

Number of Factors. An inspection of the eigenvalues resulting from the principal components analyses (Figure 2) showed that the BRC clearly measures two factors pre and post at grades K, 3, and 5.

The loadings of the items on the first two unrotated principal components (Figures 3 and 4) were very similar across time (October to April) and grade. It seems clear that the BRC measures the same two factors at each grade level.

Naming the Factors. Because the BRC appears to measure the same factors at each grade level, all grades were combined for a principal axes solution with orthogonal and oblique rotations. The oblique rotation produced the most meaningful results. The oblique pattern matrix is

presented in Figure 5. After inspection of the pattern of loadings, the first factor was named the Ready, Willing, and Able (RWA) factor because it would appear to distinguish those students who are ready, willing, and able to take part in classroom activities from those who are not. The second factor clearly measures Disruptive Classroom Behavior (DCB).

Forming Scales. Further inspection of the loadings showed that items 2, 5, 9, and 10 had much higher loadings on the RWA factor than on DCB. Items 4, 6, and 7 most clearly defined the DCB factor. Items 1, 3, and 8 loaded at a moderate level on both factors.

Items 1, 3, and 8 could have been added to either scale or both or discarded since they had moderate loadings on each factor. The decision of which scales to place them on was based on the effect the items had on the internal consistency and test-retest reliabilities of the scales. Therefore, a long and short version of each scale was created and scored. The content of each scale is listed below.

Reading, Willing, and Able

Short = 2 + 5 + 9 + 10

Long = 1 (R) + 2 + 3 (R) + 5 + 8 + 9 + 10

Disruptive Classroom Behavior

Short = 4 + 6 + 7

Long = 1 + 3 + 4 + 6 + 7 + 8 (R)

The R indicates that scoring was reversed for these items.

Internal consistency and test-retest (October to April) reliability coefficients are found in Figure 6. Figure 6 shows that both reliability estimates are higher for the long version of the RWA scale than for the short. The long version of the DCB scale was less internally consistent and only slightly more stable than the short version. The decision was made to use the long version of the RWA scale and short version of the

DCB scale in further analyses. The decision also had the advantage of having each item scored as part of one and only one scale.

Study II

The second study attempted to assess the validity of the BRC scales by showing that two select groups of students differed in their ratings on the scales.

Subjects

In the past, Title I Counselors have complained that the evaluation of the Counseling and Guidance Component in obtaining BRC ratings on a random sample of the students in their schools did not sufficiently concentrate its attention on the students with whom they worked the most. So as part of the current evaluation, counselors in the Title I schools were asked to designate twenty students with "specific behavior problems," i.e. students with whom they would work most closely. These students and a random sample of other students in the same schools were rated by their classroom teachers in October, 1978. Altogether 2,956 students were rated: 416 students with specific behavior problems and 2,540 others.

Analyses

The scale scores for the two groups were compared using the Program ANOVAR of the EDSTAT-V package (Veldman, 1971) on the AISD IBM computer.

Results

Figure 7 shows that the two groups showed marked differences in their mean scores on the two scales. The group identified as having specific behavior problems was less ready, willing, and able to participate in class than the others. They were also more disruptive than those not

so designated. Figures 7 and 8 use modified box and whiskers displays to graphically illustrate the differences in the distribution of scores for the two groups. Figure 8 shows that disruptive behavior was probably the factor which was more important in the selection of students with problems. The results show that the BRC is probably a valid measure of classroom behavior.

The differences between the two groups might have been more dramatic had the selection of students with specific problems not been so closely tied to the evaluation of the Counseling Component. The major problem was that the counselors were limited to choosing twenty students; others with low RWA ratings and high DCB ratings were included in the comparison group. In addition, a talk with a Title I counselor revealed a number of other reasons that the two distributions showed substantial overlap.

1. Counselors would tend not to select students for whom they saw little chance of improvement.
2. Students expected to move were excluded.
3. Students with problems might be concentrated at one grade level, but the counselor cannot concentrate on one grade. Therefore, some students with greater than average problems were not selected.
4. Students were not selected if there was little expectation that the classroom teacher would be willing to work with the counselor in helping the student.
5. A counselor might not select a student whose personality clashed with the counselor's.
6. A new counselor might not know which students to select.

Conclusion

In conclusion, the BRC has the qualities specified in the title of this paper. It is a practical, economical, reliable, and valid measure of student classroom behavior.

Practical because the number of items is small, and the BRC does not ask teachers to rate students on behaviors not readily observable in the classroom.

Economical because the BRC has only 10 items. Teachers can rate a student very quickly. The cost in teacher time is very low.

Reliable because the items are internally consistent. The scales show adequate constancy over six months for variables upon which change is expected.

Valid because groups expected to differ on the two scales were found to differ substantially despite difficulties with the subjective selection procedure.

Figure 1: THE NUMBER OF STUDENTS RECEIVING RATINGS ON THE BRC, 1977-78.

Grade	Number
K	295
1	352
2	329
3	284
4	255
5	247
Total	1762

Figure 2: EIGENVALUES FOR DETERMINING THE NUMBER OF FACTORS.

Factor	Grade K		Grade 3		Grade 5		All Grades	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
1	5.49	5.67	5.22	5.46	5.14	5.67	5.15	5.23
2	1.69	1.52	1.80	1.64	1.84	1.42	1.78	1.68
3	.71	.76	.76	.75	.88	.74	.73	.75
4	.59	.49	.65	.68	.53	.51	.66	.66
5	.46	.48	.45	.39	.42	.44	.46	.43
6	.33	.32	.31	.36	.35	.35	.36	.39
7	.27	.30	.28	.26	.31	.29	.31	.29
8	.21	.20	.23	.19	.20	.26	.23	.25
9	.17	.17	.18	.17	.18	.20	.21	.20
10	.08	.08	.12	.10	.14	.12	.11	.12
N	286	280	284	283	242	243	1,724	1,736

Figure 3: ITEM LOADINGS ON THE FIRST UNROTATED PRINCIPAL COMPONENT BY TIME OF RATING AND GRADE.

Item	Grade K		Grade 3		Grade 5	
	Pre	Post	Pre	Post	Pre	Post
1	.80	.80	.70	.76	.82	.81
2	-.78	-.72	-.73	-.77	-.64	-.81
3	.74	.79	.75	.80	.82	.81
4	.75	.80	.74	.77	.72	.77
5	-.47	-.38	-.34	-.31	-.33	-.40
6	.73	.80	.74	.80	.69	.75
7	.75	.78	.75	.82	.70	.79
8	-.72	-.77	-.81	-.78	-.67	-.79
9	-.77	-.74	-.75	-.68	-.78	-.76
10	-.83	-.85	-.79	-.77	-.86	-.75
N	286	280	284	283	242	241

Figure 4: ITEM LOADINGS ON THE SECOND UNROTATED PRINCIPAL COMPONENT BY TIME OF RATING AND GRADE.

Item	Grade K		Grade 3		Grade 5	
	Pre	Post	Pre	Post	Pre	Post
1	-.11	-.17	.17	.20	-.16	-.03
2	.33	.33	.43	.41	.52	.33
3	-.07	-.14	.00	-.01	-.11	-.08
4	.53	.46	.53	.48	.50	.46
5	.61	.62	.58	.63	.48	.59
6	.58	.49	.52	.41	.61	.49
7	.56	.50	.54	.44	.62	.48
8	.04	-.12	.11	.05	-.12	-.01
9	.43	.46	.48	.59	.44	.46
10	.24	.23	.41	.33	.29	.23
N	286	280	284	283	242	241

Figure 5: OBLIQUE FACTOR PATTERN LOADINGS FOR BRC ITEMS.

Item	Factor 1*		Factor 2	
	Pre	Post	Pre	Post
1. Student demands extra time from the teacher for help.	-.52	-.41	.30	.41
2. Student understands and follows directions.	.76	.78	.01	-.04
3. Student quits or gives up on assignments before completion.	-.54	-.49	.32	.39
4. Student bothers others while they are working.	-.07	-.03	.83	.87
5. Student brings things to class, initiates discussions, shows imagination.	.55	.54	.16	.14
6. Student breaks classroom or school rules.	.01	.03	.92	.92
7. Student must be reprimanded during class.	.02	.04	.94	.95
8. Student does what the teacher asks without complaint or delay.	.54	.49	-.24	-.28
9. Student is prepared and able to participate in class activities, lessons, discussions, etc.	.89	.87	.05	.04
10. Student completes work on time and in good order.	.84	.77	-.09	-.14

*Signs of loading reversed to aid interpretation.

Figure 6: INTERNAL CONSISTENCY AND TEST-RETEST RELIABILITY ESTIMATES FOR TWO BRC SCALES.

	<u>Short Version</u>		<u>Long Version</u>	
	Internal Consistency*	Test- Retest	Internal Consistency*	Test- Retest
Ready, Willing & Able	.83	.66	<u>.87</u>	<u>.71</u>
Disruptive Behavior	<u>.94</u>	<u>.70</u>	.88	.73

*Average of pre and post values.

Figure 7: COMPARISON OF STUDENTS WITH SPECIFIC BEHAVIOR PROBLEMS WITH A RANDOM SAMPLE OF OTHERS ON BRC SCALES.

Group	N	Mean	Standard Deviation	t	df	p
Ready, Willing, and Able						
Students with Specific Problems	416	30.01	12.13			
				18.45	2,954	<.0000
Others	2,540	42.42	12.81			
Disruptive Classroom Behavior						
Students with Specific Problems	416	20.75	6.59			
				25.29	2,956	<.0000
Others	2,540	11.18	7.25			

Figure 8: MODIFIED BOX AND WHISKERS COMPARING TWO GROUPS ON BRC READY, WILLING, AND ABLE SCALE.

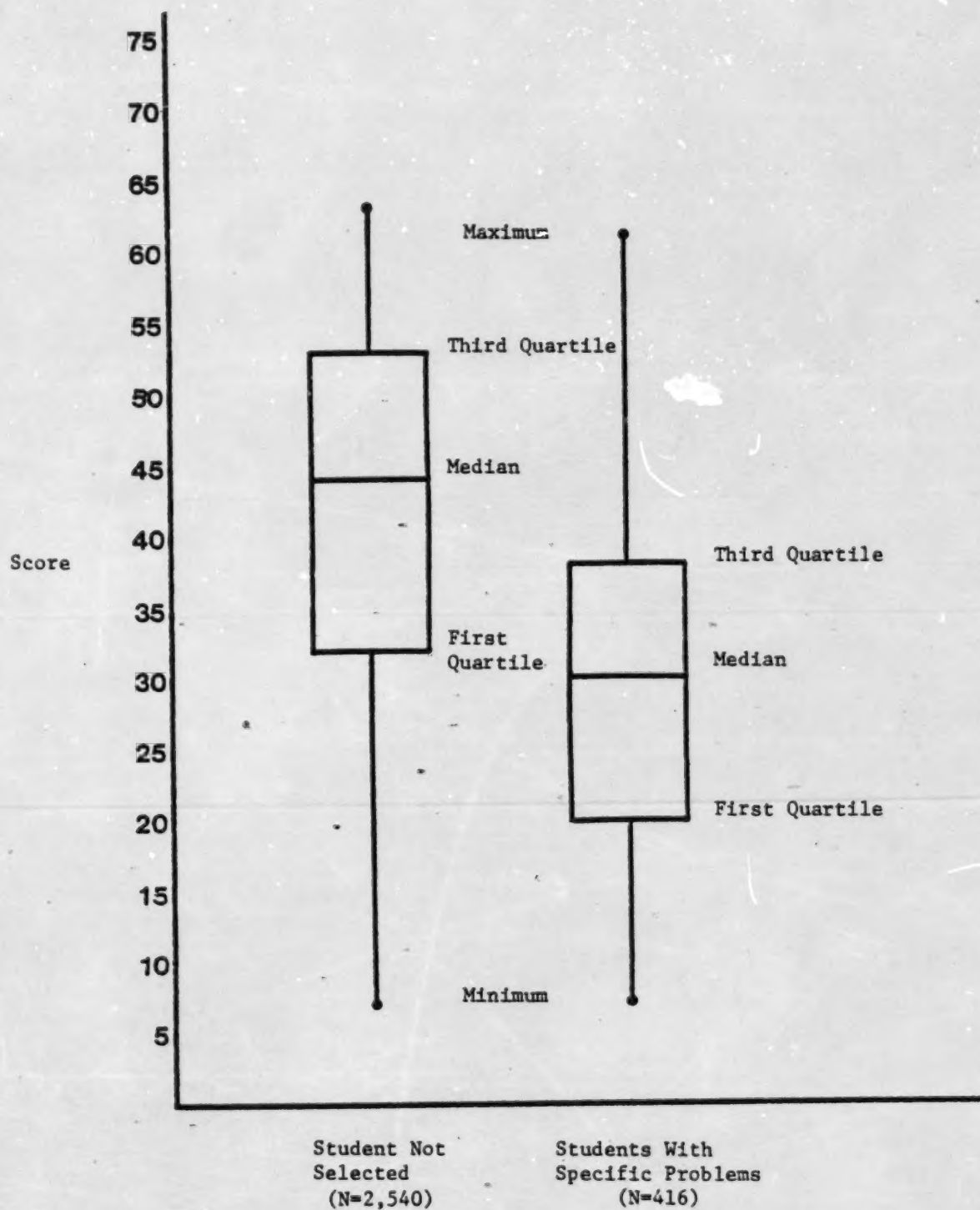
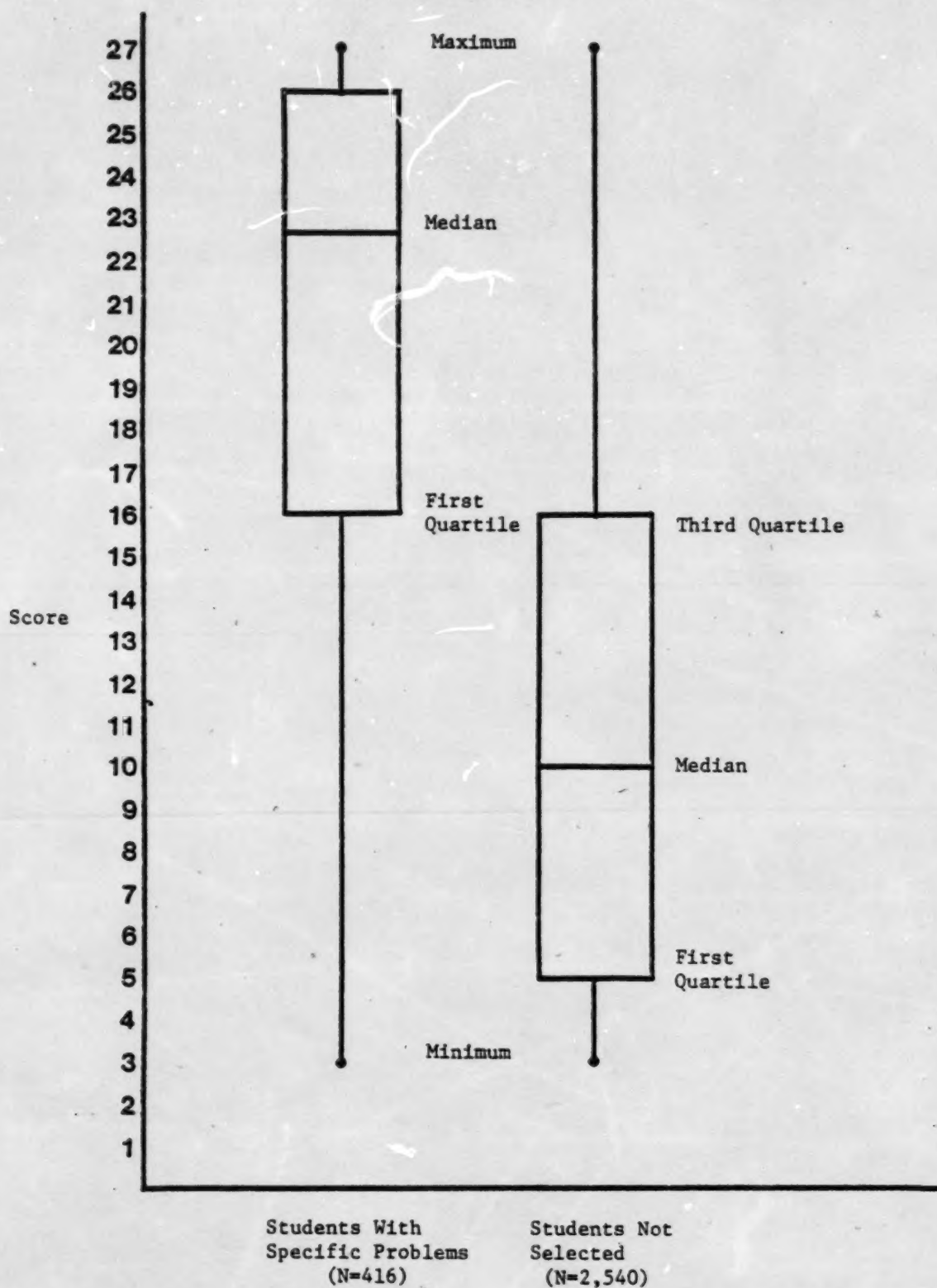


Figure 9: MODIFIED BOX AND WHISKERS COMPARING TWO GROUPS ON BRC DISRUPTIVE CLASSROOM BEHAVIOR SCALE.



References

- Buros, O. K. The seventh mental measurement yearbook. Highland Park, N. J.: Gryphon Press, 1972.
- Spivack, G. & Swift, M. Devereux elementary school behavior rating scale. Devon, Pa.: Devereux Foundation Press, 1967.
- Veldman, D. J. EDSTAT-V user's manual. Austin, Tx.: Ibid Inc., 1971.